DOCUMENT RESUME

ED 466 689 TM 034 267

AUTHOR	McMillan, James H.; Lawson, Sonya; Lewis, Kristi; Snyder, Angela
TITLE	Reporting Effect Size: The Road Less Traveled.
PUB DATE	2002-04-00
NOTE	18p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002).
PUB TYPE	Information Analyses (070) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Educational Research; *Effect Size; *Research Reports; Scholarly Journals; *Statistical Analysis

ABSTRACT

This paper is intended to raise awareness of the need to understand and include effect size statistics in research reports. The paper summarizes what is meant by effect size and why it is important in educational research. It then presents findings of a review of recent issues of four journals to determine the use of effect size by researchers. The paper also suggests some reasons effect size is not reported more commonly and offers some suggestions for increasing the use of such statistics. The four journals on which the review is based were chosen to provide a representative sample of educational research literature with a focus on quantitative studies. Of the 508 articles classified as quantitative or mixed method, 148 mentioned or calculated effect size, but only 82 articles included a calculation of effect size and at least limited discussion of magnitude or practical significance. Only 30 articles contained a calculated effect size and extensive discussion about effect size or magnitude. Most often, conclusions in these journals are made solely from the results of statistical inference tests. (Contains 4 tables and 36 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 466 689

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. McMillan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Reporting Effect Size: The Road Less Traveled

James H. McMillan
Sonya Lawson
Kristi Lewis
Angela Snyder
Virginia Commonwealth University

TM034267

Paper presented at the 2002 Annual Meeting of the American Educational Research Association, April 5, New Orleans.

BEST COPY AVAILABLE

Despite the admonitions of many statisticians, and to a lesser extent educational researchers, reporting effect size in studies other than meta-analyses remains the exception rather than the rule for educational journals (Vacha-Haase, Nilsson, Reetz, Lance, & Thompson, 2000; Thompson, 2002). Using specific effect size statistics, or even the concept of magnitude of findings as different from statistical significance, is clearly not yet integral to conducting and reporting educational research. This appears to be true even though effect size has been addressed in most statistical methods textbooks for two decades (Huberty, 2001), well-known and respected methodologists have written about effect size in leading journals (e.g., Cohen, 1994; Glass, 1976; Kirk, 1996; 2001; Rosenthal, 1991; Rosenthal & Rubin, 1982; Wilkinson & APA Task Force on Statistical Inference, 1999), and 19 journals have adopted editorial policies that require effect size reporting (Snyder, 2001; Thompson, 2002). (Notably absent from this list is the *American Educational Research Journal, Educational Researcher, Journal of Educational Psychology,* and the *Journal of Educational Research.)* Furthermore, data analysts have argued for more than three decades that measures of effect size should be reported in addition to tests of statistical significance (Olejnik & Algina, 2000).

Our intent in this article is to raise awareness for all researchers of the need to understand and include effect size statistics, not just those whose specialties are statistics or research methods. To accomplish this, we have first summarized the concept of effect size and why it is important in educational research. We then present findings from a review of recent issues of four journals to ascertain the use of effect size by researchers. The purpose of this review is to provide a more recent and larger sample of journals than has previously been reported to determine the extent to which effect size is reported and how it is used in interpreting findings. Finally, we suggest some reasons for the continued lack of use of effect size and offer some recommendations for increasing the use of these statistics and concepts.

What is Effect Size?

Effect size is a concept that refers to any of several measures of the magnitude, importance, or practicality of a difference or relationship. Some researchers, however, restrict the use of "effect size" to refer to a more specific standardized mean difference. Others prefer to use the terms "effect magnitude" or "magnitude effect," or even "magnitude-of-effect." It is argued that use of the term "magnitude" makes most sense since this more logically includes

both difference and relationship indices. However, "effect size" seems to be the preferred terminology, at least as judged by the number and percentage of writers that use these words rather than something that includes magnitude, and is consistent with the language of the recent APA Task Force on Statistical Inference (Wilkinson and APA Task Force on Statistical Inference, 1999). So, while we will use the term "effect size," others may prefer some variant of "magnitude." Regardless of the terms used, though, it is important to understand that there are many specific indices that address the issue of practicality and importance.

Following the lead of Kirk (1996) and Thompson (2002), it is useful to categorize effect size indices into three types or families, those that examine differences between two groups, those that examine strength of association, and other measures (It should be noted that Huberty (2001) argues for an additional category, group overlap). Table 1 shows a summary of some of the different indices for each of the three groups. A more complete listing can be found in Kirk (1996), and Snyder and Lawson (1993).

[Insert Table 1]

The two most common types of effect size indices are those based on a difference between two groups and those based on variance accounted-for. Cohen's *d* has become the measure of choice to examine the practicality of differences between two sets of scores. In its most general form, *d* represents differences between two sets of scores as a function of variance. Typically, *d* is calculated by dividing the difference between two means by a measure of standard deviation. Thus, *d* indicates degree of difference in terms of standard deviation units. Cohen's *d* is easily computed. Cohen's (1988) clear guidelines for interpretation (.2 *d* is "small," .5 *d* is "medium," and .8 *d* is "large") help researchers make the transition from statistical to practical significance, though, as pointed out by Snyder and Lawson (1993), such guidelines are arbitrary and should be adjusted depending on context. Measures of association are indices that indicate variance accounted for, such as phi^2, r^2, r^2_{pb}, R^2, omega^2 and eta^2. These indices are appropriate when there are two or more variables in a correlational design, or more than two groups (Rosenthal, Rosnow, & Rubin, 2000; Rosnow, Rosenthal, & Rubin, 2000). Some researchers would include unsquared measures of relationship as effect size indices (Elmore & Rotou, 2001; Kirk, 1976; Rosnow, Rosenthal, and Rubin, 2000). We have not included these measures because of the more common "degree of relationship" meaning associated with them,

even though many use terminology such as "magnitude and direction" to describe correlations. This is consistent with the 1994 *APA Publication Manual* (APA, 1994). However, it should be noted that some researchers believe squared correlation indices underestimate the magnitude of relationships (McCartney and Rosenthal; Rosnow, Rosenthal, & Rubin, 2000).

Effect Size Reporting in Journals

Ten previous reviews of different educational and psychological journals published from 1990 through 1998 indicate considerable variability between journals and little increase over time in the percentages of articles that included effect size indices (Vacha-Haase et al., 2000). Table 2 is a summary of the studies indicating the 23 journals reviewed and the percentages of effect size reported for the journal(s) included in each review. The typical methodology was to review every article in all journal issues of a designated volume. Only two of the studies reviewed four or more journals (Keselman, Huberty, Lix, Olejnik, Cribbie, Donahue, Kowalchuk, Lowman, Petoskey, Keselman, & Levin, 1998; Kirk, 1996), and a number of important journals for educational research are either omitted or reviewed for only one or two years. For example, the *Journal of Educational Psychology* and *AERJ* were reviewed only for 1994 and 1995. Keselman et al. (1998) reviewed *AERJ*, but that review was limited to 1994-95, to articles that examined between-subjects analyses. Furthermore, many of the studies were simple counts of the frequency of effect size indices used (Henson & Smith, 2000). A more meaningful review would include a determination of whether there was any discussion or interpretation of effect size.

[Insert Table 2]

In a review of a subset of five of these ten studies (Keselman et al., 1998; Snyder & Thompson, 1998; Thompson & Snyder, 1997; and Vacha-Haase & Nilsson, 1998), Henson and Smith (2000), claim "current trends suggest a slow but decisive movement toward reporting effect size measures" (p.289). However, this assertion is based on only two journals published between 1990 and 1996. Other trend data are difficult to summarize because different studies have reviewed different journals, often with unique criteria for defining "effect size." For instance, the Keselman et al. (1998) study, while reviewing articles in a large number of journals, reported effect size indices only for between-subjects statistical analyses such as ANOVA and MANOVA. Henson and Smith (2000) found that of the 23 journals reviewed in the five studies,

211 of 927 articles included at least one effect size measure (23%). Interpretation of effect size was rare, and most of the effect size indicators were measures of association (e.g., in Kirk's review (1996), 60% of the effect size indices were R^2 or the coefficient of determination). Thompson (1999b), in a review of studies examining effect size reporting, maintains that there has been very little, if any, change during the past decade.

Previous investigations of effect size reporting suggest that, despite recommendations in the 1994 *APA Publication Manual* (APA, 1994) that encourage effect size reporting, effect size measures are used in only a small percentage of studies. Moreover, only a few of these articles discuss or interpret effect size as an indication of practical significance. More recently, the APA Task Force on Statistical Inference (TFSI) (Wilkinson & TFSI, 1999), has placed greater emphasis on reporting effect size: "Always present effect sizes for primary outcomes ... We must stress again that reporting and interpreting effect sizes in the context of previously reported effects is essential to good research." (p. 599). In the new *Publication Manual of the American Psychological Association* (2001), however, the wording about effect size is strong but not completely consistent with the Task Force, "For the reader to fully understand the importance of your findings, it is almost always necessary to include some index of effect size or strength of relationship in your Results section" (p. 25). Furthermore, only 19 journals have editorial policies "requiring" effect size reporting and interpretation (see Table 3). Both Snyder (2001) and Thompson (1999c) believe that mere encouragement to report effect size has generally been ineffective.

[Insert Table 3]

Review of Trends in Four Journals

Our purpose for this review was to examine articles published in four journals, over a recent four-year period, to determine the degree and nature of effect size reporting and interpretation. We selected the years 1997-2000 to allow an analysis of whether there is a trend toward more effect size reporting. Our selection of journals was made to provide a representative sample that would provide a reasonable indication of the extent to which educational researchers used effect size indices. It included journals that are well-known, widely distributed, and publish mostly quantitative studies.

A review of each article in the four journals was conducted by one or more of the four authors. Each article was classified as quantitative, qualitative, simulation, literature review/commentary, or mixed methodology. There was also a determination of whether there was any calculation or mention of effect size, and if so, the degree to which there was discussion of what effect size meant in the context of the research. For articles that included an effect size indicee the type of estimate used was determined, using difference and relationship as two general categories of approaches. Interrater reliability was established on a subset of two journals over two years. Two independent ratings concerning whether effect size was reported, the nature of the effect size indicee, and the extent of discussion of effect, were analyzed, resulting in 95% agreement of the classifications for 50 articles.

Table 4 is a summary of the types of studies reviewed (quantitative or qualitative empirical, simulations, mixed-methods, or literature review/commentary), the extent of discussion of effect size, and methods of effect size calculation (difference or association). Of the 587 articles reviewed, 496 (85%) were classified as either quantitative or mixed methods. Of the 508 articles that were either quantitative, mixed, or simulation, 148 (29%) at least mentioned or calculated effect size. Only 82 of these 508 articles (16%) included both a calculation of effect size and at least limited discussion of magnitude or practical significance. Only 30 of 508 articles (6%) included both a calculated effect size and what we judged to be extensive discussion (typically several sentences or more of interpretation of magnitude or practical significance); about one-quarter of those (8) appeared in year 2000 volumes of *Contemporary Educational Psychology*. There were few differences between the journals, though *Contemporary Educational Psychology*, compared to the other three journals, had a higher overall percentage of articles that included calculated effect sizes and discussion (24%).

An analysis of the number of articles calculating and discussing effect size over time showed increased use from 1997 to 1999, but this trend was reversed in 2000. The percentage of articles that included calculation of effect size, with at least some discussion, for the four years from 1997 to 2000, was 20% (28 of 139), 24% (31 of 128) , 38% (44 of 115), and 37% (46 of 126), respectively. Looking at two year blocks, there was more use and discussion of effect size in 1999-2000 than in 1997-98. These findings are skewed to some extent by the large number of articles in the *Journal of Educational Psychology*, and by articles published by *Contemporary*

Educational Psychology in 2000. The *Journal of Educational Psychology*, which accounted for 47% of the quantitative or mixed method studies, showed increases in the percentages of quantitative and mixed method studies reporting effect size indices from 13% in 1998 to 21% in 1999 and 42% in 2000. The *Journal of Educational Research* showed increasing percentages of articles reporting effect size from 1997 to 1999 (22%, 39%, and 48%, respectively), but showed a decline in the percentage of articles reporting effect size to only 17% in 2000.

With respect to the methods of effect size calculation used or discussed, 43% of the total number of individual calculations (162) were difference indices. While numbers of specific types of calculations are not reported in this summary, it was clear that R^2 was clearly the most used association statistic, followed by r^2 and eta^2, while Cohen's *d* was the most common difference statistic used.

Our review of articles in these four journals supports what has been found in other studies of effect size reporting, and further confirms that there is not yet widespread understanding, application, and acceptance of such indices by those conducting, reviewing, and publishing quantitative educational research. While there is some indication that use of effect size is increasing, the trends are hopeful but not altogether clear, especially when analyzing published articles for more than simple reporting of effect size statistics. At best, some researchers include effect size calculations, but typically there is very little discussion of how effect size results should be interpreted. More often, conclusions are made solely from the results of statistical inference tests.

[Insert Table 4]

Why Isn't Effect Size Reported?

Many have speculated about the reasons for the lack of effect size reporting. As pointed out by Kirk (2001), discussing practical significance involves subjective judgment, influenced by many considerations, including the perspective of the researcher, social concerns, assessments of probable changes or differences in specific individuals or groups, costs, and the nature of the scale. Kirk goes on to suggest that the researcher has "an obligation to make this kind of judgment. No one is in a better position than the researcher who collected and analyzed the data to decide whether the effects are trivial or not." (p.214). Many researchers may be more comfortable with relatively "objective" criteria, such as statistical significance, than with more

subjective judgments. Kirk (2001) and Nickerson (2000) point out that there is widespread misunderstanding of what significance tests and *p* values tell us, and these misunderstandings contribute to an over reliance of these tests for formulating conclusions. For one, many believe that a small *p* value indicates a greater treatment effect. A common misinterpretation is concluding that a trivially small observed statistically significant difference is meaningful when sample size is large. Some may also believe that statistical significance *is* a measure of theoretical or practical significance.

Many current researchers were trained at a time when there was little acknowledgement of effect size. Without additional education or prodding by journal editors and reviewers, researchers who are not familiar with estimates of effect size will probably not use them. Moreover, McCartney and Rosenthal (2000) suggest that neither experienced researchers nor experienced statisticians have a good intuitive understanding for the practical meaning of common effect size indices. Our finding of very little discussion of practical significance, even when effect size statistics are reported, supports this contention. Another reason may be the lack of readily available software to calculate effect size. While eta-squared and R^2 may be easily determined in several programs, many of the effect size indices require either special programs or individual calculations (Olejnik & Algina, 2000). For researchers who may not be extensively trained in statistics, calculations for effect size may be perceived as too complicated.

The lack of a clear policy about effect size reporting from journals and professional organizations may contribute to a perception of the unimportance of such measures. Most journals do not currently "require" effect size reporting, and our review found that, even in those journals that do, many articles still do not include any mention or discussion of effect size. This seems to suggest a less than clear resolve on the part of reviewers and editors to enforce stated policies. Certainly a policy of "encouragement" could be interpreted to mean that it is, at best, only important, not necessary.

What Can Be Done?

The following suggestions are made with the hope that our profession can take a more proactive approach to remedy what many see as a serious deficiency in the reporting, interpretation, and use of educational research. As suggested by Kirk (2001), the approach must

be multifaceted, involving different sources, to result in the creation of new norms, expectations, and practice for reporting the practical significance of statistical findings.

1. Perhaps most importantly, AERA needs to develop and adopt policy concerning the need to report and interpret effect size indices. Such a policy would provide all educational researchers with a common understanding of what is needed and why it is needed. It would signal to all AERA members that the concept of practical significance must be addressed in reporting results from quantitative studies.
2. Since many researchers have not received formal training in reporting effect size or practical significance, AERA and other professional associations need to sponsor the development of training modules that can be used for professional development. Such modules could be web-based, print-based, and offered as training during annual meetings. It will be important in such training to focus on context-dependency issues associated with effect size estimates and on limitations to rules of thumb such as those established by Cohen (1988) (Snyder, 2001).
3. Authors of both statistics and research textbooks can lead best practice by including effect size and practical significance as major topics in their books (Hyde, 2001; Kirk, 2001; Vacha-Haase et al., 2000). Such an emphasis would effectively reach students and instructors of research methods courses.
4. Research and statistics professors need to continue to present papers and write articles stressing the importance of effect size. Papers focused on effective instruction of effect size and practical significance would be particularly helpful to those teaching the concepts. There is also need to write about effect size in journals that do not report empirical studies and are widely read by teachers and administrators.
5. Journals are the gatekeepers of quality research, and, as such, play a significant role in legitimizing the reporting and interpretation of effect size. Not only do journal policies need to insist on reporting effect size, (unless, perhaps, as suggested by Kirk (2000), authors can provide a reasonable justification for why effect size should not be reported) editors need to be diligent in requiring authors to address it. It will be most helpful if editors can suggest articles and other materials to provide authors with guidance about what is needed. Manuscript reviewers also need to insist on the

inclusion of appropriate effect size statistics as well as discussion of practical significance.

6. Statistical software packages need to include appropriate types of effect size indices that could be calculated following tests of statistical significance, much in the same way options for post hoc tests are provided. The most effective way to influence software authors and companies may be for editors and association officials to request the availability of specific effect size procedures.

One of the continuing challenges in educational research is to draw conclusions from empirical studies that will have clear implications for practice. Effect size measures provide a tool to help researchers identify what is of practical as well as statistical significance. While some data suggest that more effect size indices are being used, much more can and should be done. We believe implementation of these six suggestions would be effective in further raising awareness and understanding, and would increase proper use of effect size estimates and the concept of practical significance in ways that would enhance the overall validity and use of research findings.

References

American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2ed). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Cohen, J. (1994). The earth is round ($p<.05$). *American Psychologist, 49*, 997-1003.

Elmore, P. B., & Rotou, O. (2001). A primer on basic effect size concepts. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle.

Fleiss, J. L. (1994). Measures of effect size for categorical data. In H. Cooper & L. V, Hedges (Eds.), *The handbook of research synthesis* (pp. 245 - 260). New York: Russell Sage Foundation.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3-8.

Hedges, L. V. (1981). Distributional theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128.

Henson, R. K., & Smith, A. D. (2000). State of the art in statistical significance and effect size reporting: A review of the APA Task force report and current trends. *Journal of Research and Development in Education, 33*, 285-296.

Huberty, C. J. (2001). A history of effect size indices. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle.

Hyde, J. S. (2001). Reporting effect sizes: The roles of editors, textbook authors, and publication manuals. *Educational and Psychological Measurement, 61*, 225-228.

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., Kowalchuk, R. K., Lowman, L. L., Petoskey, M. D., Keselman, J. C., & Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA Analyses. *Review of Educational Research 68 (3),* 350-386.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56*, 746-759.

Kirk, R. E. (2001). Promoting good statistical practices: Some suggestions. *Educational and Psychological Measurement, 61*, 213-218.

Lance, T., & Vacha-Haase, T. (1998). The counseling psychologist: Trends and usages of statistical significance testing. Paper presented at the Annual Meeting of the American Psychological Association, San Francisco.

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71 (1),* 173-180.

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods, 5*, 241-301.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology 25*, 241-286.

Preece, P. F. W. (1983). A measure of experimental effect size based on success rates. *Educational and Psychological Measurement, 43*, 763-766.

Rosenthal, R. (1991). Effect sizes: Pearson's correlation, its display via the BESD, and alternative indices. *American Psychologist, 46*, 1086-1087.

Rosenthal, R., Rosnow, R. I., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach.* NY: Cambridge University Press.

Rosenthal, R., & Rubin, D.B. (1982). A simple general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*, 166-169.

Rosnow, R.I., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological Science, 11 (6)*, 446-453.

Snyder, P. A. (2001). Treatment of effect indices in journal editorial policies: An editor's perspective. Paper presented at the 2001 Annual Meeting of the American Educational Research Association.

Snyder, P. A., & Lawson, S. (1993). Effect size estimates. *Journal of Experimental Education, 61*, 293-316.

Snyder, P. A., & Thompson, B. (1998). Use of tests of statistical significance and other analytic choices in a school psychology journal: Review of practices and suggested alternatives. *School Psychology Quarterly, 13*, 335-348.

Thompson, B. (1999a). Common methodology mistakes in education research, revisited, along with a primer on both effect sizes and the bootstrap. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal.

Thompson, B. (1999b). Improving research clarity and usefulness with effect size indices as supplements to statistical significance tests. *Exceptional Children, 65(3)*, 329-337.

Thompson, B. (1999c). Journal editorial policies regarding statistical significance: Heat is to fire as p is to importance. *Educational Psychology Review, 11*, 157-169.

Thompson, B. (2002). "Statistical," 'practical," and "clinical": How many kinds of significance do counselors need to Consider? *Journal of Counseling and Development.80*, 64-71.

Thompson, B., & Snyder, P. A. (1997). Statistical significance testing practices in the *Journal of Experimental Education*. *Journal of Experimental Education, 66, 75*-83.

Thompson, B., & Snyder, P. A. (1998). Statistical significance and reliability analyses in recent JCD research articles. *Journal of Counseling and Development, 76*, 436-441.

Vacha-Haase, T., & Nilsson, J. E. (1998). Statistical significance reporting: Current trends and usages with *MECD*. *Measurement and Evaluation in Counseling and Development, 31*, 46-57.

Vacha-Haase, T., & Ness, C. (1999). Statistical significance testing as it relates to practice: Use within *Professional Psychology*. *Professional Psychology Research and Practice, 30*, 104-105.

Vacha-Haase, T., Nilsson, J. E., Reetz, D. R., Lance, T. S., & Thompson, B. (2000). Reporting practices and APA editorial policies regarding statistical significance and effect size. *Theory & Psychology, 10(3)*, 413-425.

Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604.

Table 1

Effect Size Measures

Measures of Magnitude of Difference	Measures of Strength of Association	Other Measures
Cohen's (1988) *d*	phi^2	Rosenthal and Rubin's (1982)
Glass's (1976) g for meta-analysis	r^2	binomial effect size
Hedges's (1981) g	R^2	Fleiss's (1994) categorical data effect size
	Eta^2	
	$Omega^2$	Preece's (1983) ratio of success rates
		Relative risk
		Risk difference

Table 2[1]

Previous Effect Size Reporting Practices

Study	Journal(s)	Year(s)	Effect Size Reported (%)
Keselman et al. (1998)	*American Educational Research Journal* *Child Development* *Cognition and Instruction* *Contemporary Educational Psychology* *Developmental Psychology* *Educational Technology, Research and Development* *Journal of Applied Psychology* *Journal of Counseling Psychology* *Journal of Educational Computing Technology* *Journal of Educational Psychology* *Journal of Experimental Child Psychology* *Sociology of Education*	1994-95	10 (average)
Kirk (1996)	*Journal of Applied Psychology*	1995	23
	Journal of Educational Psychology	1995	45
	Journal of Experimental Psychology	1995	88
	Journal of Personality and Social Psychology	1995	53
Lance & Vacha-Haase (1998)	*The Counseling Psychologist*	1995-(	41
Vacha-Haase et al. (2000)	*Journal of Counseling Psychology*	1995-7	53
	Psychology and Aging	1995-7	47
Snyder & Thompson (1998)	*School Psychology Quarterly*	1990-6	54
Thompson (1999)	*Exceptional Children*	1996-8	13
Thompson & Snyder (1997)	*Journal of Experimental Education*	1994-7	36
Thompson & Snyder (1998)	*Journal of Counseling and Development*	1996	10
Vacha-Haase & Ness (1999)	*Professional Psychology: Research and Practice*	1995-7	21
Vacha-Haase & Nilsson (1998)	*Measurement & Evaluation in Counseling and Development*	1990-6	35

[1] Table was adapted from Vache-Haase et al. (2000).

Table 3

Journals with Editorial Policies Requiring Effect Size Reporting and Interpretation

Career Development Quarterly
Contemporary Educational Psychology
Exceptional Children
Educational and Psychological Measurement
Exceptional Children
Journal of Agricultural Education
Journal of Applied Psychology
Journal of Community Psychology
Journal of Consulting and Clinical Psychology
Journal of Counseling & Development
Journal of Early Intervention
Journal of Educational and Psychological Consultation
Journal of Experimental Education
Journal of Learning Disabilities
Language Learning
Measurement and Evaluation in Counseling and Development
The Professional Educator
Reading and Writing
Research in the Schools

Table 4

Frequencies of Different Types of Studies and Effect Magnitude Measures Used in Four Journals

Journal, Year, and Total Number of	Type of Study			Reporting and Discussion of Magnitude[1]				Methods of Calculation[2]	
Articles	Quantitative	Qualitative	Other[3]	None	1	2	3	Association	Difference
Journal of Educational Psychology									
1997 (n=64)	59	1	M = 2 C/LR = 2	49	7	4	0	17	12
1998 (n=58)	54	2	M = 2 C/LR = 0	49	4	3	0	9	9
1999 (n=64)	60	0	M = 1 C/LR = 3	48	9	2	2	12	9
2000 (n=70)	67	2	M = 0 C/LR = 1	39	13	11	4	14	11
Journal of Experimental Education									
1997 (n=26)	16	0	S = 5 LR/C = 5	11	2	3	0	3	2

[1]none = no calculation or discussion, 1= either calculated or discussed, but not both, 2= calculation with limited discussion, 3= calculation with extensive discussion

[2] More than one method could be used in a single article. Measures of association include r^2, R^2, omega^2; effect size measures include Cohen's *d* and *f*, standardized effect size, and eta^2,

[3] C/LR=Commentary/Literature Review, M=Mixed quantitative/qualitative, S=Simulation

1998 (n=23)	14	0	S = 2 LR/C = 7	10	2	0	2	1	1
1999 (n=21)	13	0	S = 3 LR/C = 5	1	8	4	0	7	6
2000 (n=22)	14	0	M = 2 S = 2 LR/C = 4	10	2	2	0	1	3
Journal of Educational Research									
1997 (n=40)	33	4	M=3	25	5	1	2	6	2
1998 (n=40)	33	3	M=3 LR/C = 1	19	2	5	7	9	4
1999 (n=31)	22	4	M=3 LR/C = 2	15	8	4	0	4	2
2000 (n=34)	24	7	LR/C = 3	14	2	2	0	6	4
Contemporary Educational Psychology									
1997 (n=26)	21	2	LR/C = 3	19	2	2	0	4	0
1998 (n=22)	19	1	M=1 LR/C = 1	13	0	1	5	3	0
1999 (n=19)	13	1	LR/C = 5	7	1	6	0	7	0
2000 (n=27)	16	1	M = 1 LR/C = 9	6	0	2	8	6	4

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM034267

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Reporting Effect Size: The Road Less Traveled

Author(s): James McMillan, Sonya Lawson, Kristi Lewis, Angela Snyder

Corporate Source: Virginia Commonwealth University

Publication Date: 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1 ☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A ☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B ☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: James McMillan

Printed Name/Position/Title: Dr. James McMillan, Professor

Organization/Address: Virginia Commonwealth U. Box 842020, Richmond VA 23284

Telephone: 804 828 1332 FAX: 225-3554

E-Mail Address: jmcmillan@vcu.edu Date: 5/23/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of Maryland
ERIC Clearinghouse on Assessment and Evaluation
1129 Shriver Laboratory
College Park, MD 20742
Attn: Acquisitions

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)
PREVIOUS VERSIONS OF THIS FORM ARE OBSOLETE.